RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _09/647,543_

Source: _JFW16_

Date Processed by STIC: _11/07/2005_

ENTERED

IFW16

RAW SEQUENCE LISTING   DATE: 11/07/2005
PATENT APPLICATION: US/09/647,543   TIME: 11:27:44

Input Set : **A:\Sequence.txt**
Output Set: **N:\CRF4\11072005\I647543.raw**

```
 3 <110> APPLICANT: URQUIMA, S.A.
 5 <120> TITLE OF INVENTION: Promotor and constructions for expression of
 6       recombinant proteins in filamentous fungi
 8 <130> FILE REFERENCE: Thaumatin 2 .
```
```
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/647,543
C--> 11 <141> CURRENT FILING DATE: 2000-10-02
```
```
13 <150> PRIOR APPLICATION NUMBER: ES 9800699
14 <151> PRIOR FILING DATE: 1998-04-02
16 <160> NUMBER OF SEQ ID NOS: 13
18 <170> SOFTWARE: PatentIn Ver. 2.1
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 2570
22 <212> TYPE: DNA
23 <213> ORGANISM: Aspergillus awamorii
25 <220> FEATURE:
26 <221> NAME/KEY: intron
27 <222> LOCATION: (785)..(850)
29 <220> FEATURE:
30 <221> NAME/KEY: intron
31 <222> LOCATION: (1414)..(1471)
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (741)..(785)
37 <220> FEATURE:
38 <221> NAME/KEY: CDS
39 <222> LOCATION: (850)..(1413)
41 <220> FEATURE:
42 <221> NAME/KEY: CDS
43 <222> LOCATION: (1472)..(2242)
45 <400> SEQUENCE: 1
46 tctagattgc gacggcgtat tgcttatcct tagtaggact ccctaatgga ttccgagcaa 60
48 gaaaagactg tttggcgtgt accaatggct catagtacca gcaagagaag aattttctct 120
50 ctcgcttcga gaaagcaatc aaaaaaaaat cctatcctac cctaccctac cctaatactt 180
52 ccattgccac ccgattcctc ccgatagtag agcgggcgac tgccatttgg cgggcgggcc 240
54 agcggattcc cgccgataga taacgggcag attctgtgac ctcaaactat cgactaacag 300
56 cccgaacttc ggcggccacc gccaaacccg ccccggaagc cggcctcatt tgccgtttgg 360
58 gcgtgccagg aaatgccgcc tgcagcggag actccctagt gtggtctgtg ttgcctgtgt 420
60 cgtctgtgta gtatactagt tactagtcta ctactgtaca gtggatggcc tgaggggggg 480
62 actttatgtc cgactccggc tgttctcctc cctctatcca ctctaccctc ttccctctct 540
64 tctgtctttc tccccgctct cgcccctccc ctcctcgaaa acataaatcg gcctttcccc 600
66 ctcgccatct tcttcttctt ctccctctcc tttctctttc ttcttcagac tacttctctt 660
68 tctttcatct tttctctata ttcctgtttt cctagatacc ccagttaaaa aagttctctc 720
70 aatcaatcct ccccttcaga atg tct aac ctt cct cac gag ccc gag ttc gag 773
```

RAW SEQUENCE LISTING   DATE: 11/07/2005
PATENT APPLICATION: US/09/647,543   TIME: 11:27:44

Input Set : **A:\Sequence.txt**
Output Set: **N:\CRF4\11072005\I647543.raw**

```
71                      Met Ser Asn Leu Pro His Glu Pro Glu Phe Glu
72                       1               5                   10
74 cag gcc tac aag ggtatgttcc attgcccctc cgaaattgat gatggaaaaa        825
75 Gln Ala Tyr Lys
76               15
78 aaattctaac aacatcctct taca gag ctt gcc tcg acc ctt gag aac tcc     876
79                           Glu Leu Ala Ser Thr Leu Glu Asn Ser
80                                                20
82 acc ctc ttc cag aag aac ccc gaa tac cgc aag gcc ctt gct gtc gtc    924
83 Thr Leu Phe Gln Lys Asn Pro Glu Tyr Arg Lys Ala Leu Ala Val Val
84  25               30                  35                   40
86 tcc gtc ccc gag cgt gtc atc cag ttc cgt gtc gtc tgg gag gat gat    972
87 Ser Val Pro Glu Arg Val Ile Gln Phe Arg Val Val Trp Glu Asp Asp
88                   45                  50                   55
90 gcc ggc aac gtc cag gtc aac cgc ggt ttc cgt gtc cag ttc aac agc    1020
91 Ala Gly Asn Val Gln Val Asn Arg Gly Phe Arg Val Gln Phe Asn Ser
92                   60                  65                   70
94 gcc ctc ggt ccc tac aag ggt ggt ctt cgt ttc cac ccc tcc gtc aac    1068
95 Ala Leu Gly Pro Tyr Lys Gly Gly Leu Arg Phe His Pro Ser Val Asn
96               75                  80                  85
98 ttg tcc atc ctc aag ttc ctt ggt ttc gag cag atc ttc aag aat gct    1116
99 Leu Ser Ile Leu Lys Phe Leu Gly Phe Glu Gln Ile Phe Lys Asn Ala
100          90                  95                  100
102 ctc act ggc ctg aac atg ggt ggt ggt aag ggt ggt tcc gac ttc gac   1164
103 Leu Thr Gly Leu Asn Met Gly Gly Gly Lys Gly Gly Ser Asp Phe Asp
104 105                 110                 115                 120
106 ccc aag ggc aag tcc gac aac gag atc cgt cgc ttc tgt gtt tcc ttc   1212
107 Pro Lys Gly Lys Ser Asp Asn Glu Ile Arg Arg Phe Cys Val Ser Phe
108                 125                 130                 135
110 atg acc gag ctc tgc aag cac atc ggt gcc gac act gat gtt ccc gct   1260
111 Met Thr Glu Leu Cys Lys His Ile Gly Ala Asp Thr Asp Val Pro Ala
112                 140                 145                 150
114 ggt gac atc ggt gtc acc ggt cgt gag gtc ggt ttc ctc ttc ggc cag   1308
115 Gly Asp Ile Gly Val Thr Gly Arg Glu Val Gly Phe Leu Phe Gly Gln
116          155                 160                 165
118 tac cgc aag atc cgc aac cag tgg gag ggt gtt ctc acc ggt aag ggt   1356
119 Tyr Arg Lys Ile Arg Asn Gln Trp Glu Gly Val Leu Thr Gly Lys Gly
120          170                 175                 180
123 ggc agc tgg ggt ggt tcc ctc atc cgc cct gag gcc acc ggt tac ggt   1404
124 Gly Ser Trp Gly Gly Ser Leu Ile Arg Pro Glu Ala Thr Gly Tyr Gly
125 185                 190                 195                 200
127 gtt gtc tac gtatgtcaat tcctcttctt atgattatct atgtataaca           1453
128 Val Val Tyr
130 gcgactaacg cgtaacag tac gtc gag cac atg att gct cac gcc acc aac    1504
131                     Tyr Val Glu His Met Ile Ala His Ala Thr Asn
132                         205                 210
134 ggc cag gag tcc ttc aag ggc aag cgc gtt gcc atc tcc ggt tcc ggt   1552
135 Gly Gln Glu Ser Phe Lys Gly Lys Arg Val Ala Ile Ser Gly Ser Gly
136 215                 220                 225                 230
```

```
138 aac gtt gcc cag tac gcc gcc ctc aag gtc att gag ctc ggc ggt tcc   1600
139 Asn Val Ala Gln Tyr Ala Ala Leu Lys Val Ile Glu Leu Gly Gly Ser
140             235             240             245
142 gtc gtc tcc ctg agc gac acg cag ggc tcc ctc atc atc aac ggc gag   1648
143 Val Val Ser Leu Ser Asp Thr Gln Gly Ser Leu Ile Ile Asn Gly Glu
144             250             255             260
146 ggt agc ttc acc ccc gag gag atc gag ctc atc gct cag acc aag gtc   1696
147 Gly Ser Phe Thr Pro Glu Glu Ile Glu Leu Ile Ala Gln Thr Lys Val
148             265             270             275
150 gag cgc aac gag ctc gcc agc atc gtc ggt gct gct ccc ttc agc gac   1744
151 Glu Arg Asn Glu Leu Ala Ser Ile Val Gly Ala Ala Pro Phe Ser Asp
152     280             285             290
154 gcc aac aag ttc aag tac att gct ggt gcc cgc ccc tgg gtt cac gtc   1792
155 Ala Asn Lys Phe Lys Tyr Ile Ala Gly Ala Arg Pro Trp Val His Val
156 295             300             305             310
158 ggc aag gtc gac gtc gct ctc ccc tcc gct acc cag aac gaa gtt tcc   1840
159 Gly Lys Val Asp Val Ala Leu Pro Ser Ala Thr Gln Asn Glu Val Ser
160             315             320             325
162 ggc gag gag gcc cag gtc ctc atc aac gct ggc tgc aag ttc atc gcc   1888
163 Gly Glu Glu Ala Gln Val Leu Ile Asn Ala Gly Cys Lys Phe Ile Ala
164             330             335             340
166 gag ggt tcc aac atg ggt tgc acc cag gag gcc atc gac acc ttc gag   1936
167 Glu Gly Ser Asn Met Gly Cys Thr Gln Glu Ala Ile Asp Thr Phe Glu
168             345             350             355
170 gcc cac cgt acc gcc aac gct ggc gcg gct gcc atc tgg tac gcc ccc   1984
171 Ala His Arg Thr Ala Asn Ala Gly Ala Ala Ala Ile Trp Tyr Ala Pro
172     360             365             370
174 ggt aag gcc gcc aac gcc ggt ggt gtc gct gtc tcc ggt ctg gag atg   2032
175 Gly Lys Ala Ala Asn Ala Gly Gly Val Ala Val Ser Gly Leu Glu Met
176 375             380             385             390
178 gct cag aac tct gcc cgc ctc agc tgg act tct gag gag gtt gat gcc   2080
179 Ala Gln Asn Ser Ala Arg Leu Ser Trp Thr Ser Glu Glu Val Asp Ala
180             395             400             405
184 cgt ctt aag gac atc atg cgc gac tgc ttc aag aac ggt ctt gag act   2128
185 Arg Leu Lys Asp Ile Met Arg Asp Cys Phe Lys Asn Gly Leu Glu Thr
186     410             415             420
188 gct cag gag tac gcc acc ccc gct gag ggt gtc ctg cct tcc ctg gtg   2176
189 Ala Gln Glu Tyr Ala Thr Pro Ala Glu Gly Val Leu Pro Ser Leu Val
190             425             430             435
192 acc gga tcc aac att gcc ggt ttc acc aag gtg gct gcc gcc atg aag   2224
193 Thr Gly Ser Asn Ile Ala Gly Phe Thr Lys Val Ala Ala Ala Met Lys
194     440             445             450
196 gac cag ggt gac tgg tgg taaatgcgga aagccgcaaa cccccgcggc          2272
197 Asp Gln Gly Asp Trp Trp
198 455                     460
200 ttatgtcatg acgattatgt agtttgatgt tcccttttcag cgcggatgga tagaggcgcc 2332
202 ggtgttttct tgctagttta gatggatgca taatgatatc cttttcttaa tcctcaaatt 2392
204 cttgtaattt gttgtatcaa tagtagataa tacaactgta gtcaactacc cttgcatctt 2452
206 cactatttgc agatgcattc atctctattc cgagcacatg cacaaaccca tgggaccgca 2512
```

```
208 gttcactagt acttagcctg ttatcttccc tctatcgcat cttaaacaac tatctaga    2570
211 <210> SEQ ID NO: 2
212 <211> LENGTH: 460
213 <212> TYPE: PRT
214 <213> ORGANISM: Aspergillus awamorii
216 <400> SEQUENCE: 2
217 Met Ser Asn Leu Pro His Glu Pro Glu Phe Glu Gln Ala Tyr Lys Glu
218  1               5                  10                  15
220 Leu Ala Ser Thr Leu Glu Asn Ser Thr Leu Phe Gln Lys Asn Pro Glu
221             20                  25                  30
223 Tyr Arg Lys Ala Leu Ala Val Val Ser Val Pro Glu Arg Val Ile Gln
224         35                  40                  45
226 Phe Arg Val Val Trp Glu Asp Asp Ala Gly Asn Val Gln Val Asn Arg
227     50                  55                  60
229 Gly Phe Arg Val Gln Phe Asn Ser Ala Leu Gly Pro Tyr Lys Gly Gly
230  65                  70                  75                  80
232 Leu Arg Phe His Pro Ser Val Asn Leu Ser Ile Leu Lys Phe Leu Gly
233             85                  90                  95
235 Phe Glu Gln Ile Phe Lys Asn Ala Leu Thr Gly Leu Asn Met Gly Gly
236             100                 105                 110
238 Gly Lys Gly Gly Ser Asp Phe Asp Pro Lys Gly Lys Ser Asp Asn Glu
239         115                 120                 125
241 Ile Arg Arg Phe Cys Val Ser Phe Met Thr Glu Leu Cys Lys His Ile
242     130                 135                 140
245 Gly Ala Asp Thr Asp Val Pro Ala Gly Asp Ile Gly Val Thr Gly Arg
246 145                 150                 155                 160
248 Glu Val Gly Phe Leu Phe Gly Gln Tyr Arg Lys Ile Arg Asn Gln Trp
249             165                 170                 175
251 Glu Gly Val Leu Thr Gly Lys Gly Gly Ser Trp Gly Gly Ser Leu Ile
252             180                 185                 190
254 Arg Pro Glu Ala Thr Gly Tyr Gly Val Val Tyr Tyr Val Glu His Met
255         195                 200                 205
257 Ile Ala His Ala Thr Asn Gly Gln Glu Ser Phe Lys Gly Lys Arg Val
258     210                 215                 220
260 Ala Ile Ser Gly Ser Gly Asn Val Ala Gln Tyr Ala Ala Leu Lys Val
261 225                 230                 235                 240
263 Ile Glu Leu Gly Gly Ser Val Val Ser Leu Ser Asp Thr Gln Gly Ser
264             245                 250                 255
266 Leu Ile Ile Asn Gly Glu Gly Ser Phe Thr Pro Glu Glu Ile Glu Leu
267             260                 265                 270
269 Ile Ala Gln Thr Lys Val Glu Arg Asn Glu Leu Ala Ser Ile Val Gly
270         275                 280                 285
272 Ala Ala Pro Phe Ser Asp Ala Asn Lys Phe Lys Tyr Ile Ala Gly Ala
273     290                 295                 300
275 Arg Pro Trp Val His Val Gly Lys Val Asp Val Ala Leu Pro Ser Ala
276 305                 310                 315                 320
278 Thr Gln Asn Glu Val Ser Gly Glu Glu Ala Gln Val Leu Ile Asn Ala
279             325                 330                 335
281 Gly Cys Lys Phe Ile Ala Glu Gly Ser Asn Met Gly Cys Thr Gln Glu
```

```
282             340              345              350
284 Ala Ile Asp Thr Phe Glu Ala His Arg Thr Ala Asn Ala Gly Ala Ala
285         355              360              365
287 Ala Ile Trp Tyr Ala Pro Gly Lys Ala Ala Asn Ala Gly Gly Val Ala
288     370              375              380
290 Val Ser Gly Leu Glu Met Ala Gln Asn Ser Ala Arg Leu Ser Trp Thr
291 385              390              395              400
293 Ser Glu Glu Val Asp Ala Arg Leu Lys Asp Ile Met Arg Asp Cys Phe
294         405              410              415
296 Lys Asn Gly Leu Glu Thr Ala Gln Glu Tyr Ala Thr Pro Ala Glu Gly
297         420              425              430
299 Val Leu Pro Ser Leu Val Thr Gly Ser Asn Ile Ala Gly Phe Thr Lys
300     435              440              445
302 Val Ala Ala Ala Met Lys Asp Gln Gly Asp Trp Trp
303     450              455              460
306 <210> SEQ ID NO: 3
307 <211> LENGTH: 33
308 <212> TYPE: DNA
309 <213> ORGANISM: Artificial Sequence
311 <220> FEATURE:
312 <223> OTHER INFORMATION: Description of Artificial Sequence:
313       Oligonucleotide ThS1
315 <400> SEQUENCE: 3
316 cgaatgaaaa ggaaaaggat ggccaccttc gag                          33
319 <210> SEQ ID NO: 4
320 <211> LENGTH: 18
321 <212> TYPE: DNA
322 <213> ORGANISM: Artificial Sequence
324 <220> FEATURE:
325 <223> OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
326       ThS2
328 <400> SEQUENCE: 4
329 ttattaggcg gtggggca                                          18
332 <210> SEQ ID NO: 5
333 <211> LENGTH: 18
334 <212> TYPE: DNA
335 <213> ORGANISM: Artificial Sequence
337 <220> FEATURE:
338 <223> OTHER INFORMATION: Description of Artificial Sequence:
339       Oligonucleotide IA
341 <400> SEQUENCE: 5
342 atgtctaacc ttcctcac                                          18
345 <210> SEQ ID NO: 6
346 <211> LENGTH: 18
347 <212> TYPE: DNA
348 <213> ORGANISM: Artificial Sequence
350 <220> FEATURE:
351 <223> OTHER INFORMATION: Description of Artificial Sequence:
352       Oligonucleotide IB
```

VERIFICATION SUMMARY   DATE: 11/07/2005
PATENT APPLICATION: US/09/647,543   TIME: 11:27:45

Input Set : **A:\Sequence.txt**
Output Set: **N:\CRF4\11072005\I647543.raw**

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date